



**UNIVERSITA' DEGLI STUDI DI PADOVA**  
**DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI**  
**"M.FANNO"**

**CORSO DI LAUREA MAGISTRALE / SPECIALISTICA IN**

*Economics and Finance – Economia e Finanza*

**TESI DI LAUREA**

*"Interest Rates Term Structures: the Effects of Macro Factors"*

**RELATORE:**

**CH.MO PROF. *Nunzio Cappuccio***

**LAUREANDO *Fabio Vian***

**MATRICOLA N. *1084057***

**ANNO ACCADEMICO 2015 – 2016**

Il candidato dichiara che il presente lavoro è originale e non è già stato sottoposto, in tutto o in parte, per il conseguimento di un titolo accademico in altre Università italiane o straniere. Il candidato dichiara altresì che tutti i materiali utilizzati durante la preparazione dell'elaborato sono stati indicati nel testo e nella sezione "Riferimenti bibliografici" e che le eventuali citazioni testuali sono individuabili attraverso l'esplicito richiamo alla pubblicazione originale.

Firma dello studente

---





## SUMMARY

---

1. Introduction .....	1
1.1 Motivation for Bond Yields Research .....	2
2. Asset Pricing .....	5
2.1 Time Variation of Expected Returns .....	5
2.2.1 Forward Rates .....	8
3. Pricing Bonds in Continuous Time .....	11
3.1 The Local Expectation Hypothesis .....	13
3.1.1 PDE for Bond Prices: the Role of LEH .....	14
3.2 Affine Term Structure Models .....	17
3.2.1 Affine Diffusions .....	17
4. Interest Rates Term Structure: the Model .....	21
4.1 The Macro - Finance Model .....	21
5. The Model .....	31
5.1 Yields Data .....	31
5.2 Model 1: A Single Factor No Arbitrage Model Models .....	33
5.3 Model 2: A Three Unobservable Factors No Arbitrage Model .....	39
5.4 Model 3: A No Arbitrage Model with Macro Factors .....	44
6. Conclusions .....	51
7. References .....	53
Websites .....	54



The model we will be describing all the way through this paper is based on two broad categories of terms structure models: affine term structure models and vector autoregression models. The first stream of literature shapes the general framework of the model, that will later be linked to the VAR models through some similarities that Ang and Piazzesi (2003) first noted. It is precisely for this reason that we need to produce some basic knowledge and overall description of Affine Term Structure Models (ATSMs).

The research in the field of bond yields mostly moved in the continuous time context, mainly with the aid of Partial Differential Equations, even if many other models focused on VARs and an important distinction must be made in this sense: the macroeconomics variables that were usually included in such econometric models of course described only partially bond yields movement; on the other side stochastic calculus implied in ATSMs managed to give a more complete description of bond yields movements, though losing some ground for what concerns tractability and closed solutions. In this space left out between the two approaches we may find some reason to bring them together: there are some topics that draw aside VARs from ATSM models.

The most important among such topics is that the macroeconomic variables typically used in simple VARs neglect the core features of bonds, the first one being that bonds are first of all assets. Beyond the traditional view of bonds as safer assets than equity instruments, if we put a bond in a short term trading context, risk increases along with the maturity of the bond and, as a natural consequence, the risk-averse investor will require some premium for the risk he bears. This means that to avoid arbitrage opportunities in bond yields, the expected future short rates, adjusted for such risk, should well describe long maturities yields at the present time: this is what the liaison between the different maturities at a given time is made of, this is what risk premia specifications in ATSMs and what the cross-equation restrictions in modern VAR models are intended to describe usually.

The definition of ATSMs is not univocal across the literature, as some details may differ among authors. Generally ATSMs encircles all those term structure models that include some no-

arbitrage restriction and a linear relationship between bond yields and some vector of state variables  $x$ , where the constant term and the slope coefficient are functions of maturity of a bond.

The breaking point in term structure literature came with the work of Vasicek (1977) that firstly introduced such kind of models and using a single state variable being the present short rate, whose movements were defined by a Markov process. ATSMs then grew in number and diversification starting from Vasicek-like models, with a single state variable, and evolving to multifactor models where the state variable became a vector.

The common framework of ATSM is by the way a description of yields as risk-adjusted expected future rates. This adjustment strictly relates to the definition of risk premia and excess bond returns. In the earliest models in this stream of literature a time-constant term premia on bonds, even if this led to bond yields definition that did not really well suit observed data.

## 1.1

### MOTIVATION FOR BOND YIELDS RESEARCH

---

There are different motivations to justify bond yields research, among these the first one is surely forecasting.

After Vasicek work, many other papers underlined what seems to be a rational relationship between bonds of different maturities. The fact that long maturities bond yields are usually found to drive short-term rates means that present time yield curve contains some information about the future of the economy, or at least about the expectation on the economy. As the main model will explain in next chapters, the relationship between yields and macroeconomic variables and the forward-looking feature of longer-term maturity yields may create some mechanism that should be in some way useful to get some knowledge about the future path of the economy. This should not be surprising, given the great role of expectations in economics, but as Stock and Watson (2003) may suggest, this kind of relationship may not be stable across periods.

Another motivation that may be quoted is linked to the central bank monetary policy decisions. As it is widely known, central banks have only control on shortest maturities bond yields, but internal demand usually is affected by the short end of the yield curve: most decisions by economic agents, even on a micro basis, are based on expectations about future and thus on long term yields. Again this kind of chain connecting the monetary authority and the stimulation



of the economy passes through the connection among different maturity bonds, even though some kind of precision is lost through the stages: the tentative activity of central bank usually is acting on inflation and real activity through short rates, and by intentionally communicating its own intentions, it usually acts on economy and, more concretely, on financial markets.

Finally, hedging and derivatives pricing may imply some use of yield curve. This is so since derivatives linked to interest rates may need such models to be priced, or other risks normally taken by banks may need to be hedged using yield curves and the linkages between yield curve movements and the dynamics of the state of the economy.



# 2

## ASSET PRICING

---

### 2.1

#### TIME VARIATION OF EXPECTED RETURNS

---

To have a good view on the general model we will produce it is definitely useful to get which kind of consequences assumptions on expected returns may have. If we consider a scenario having  $k$  possible states of nature having probability  $\pi ( k )$ , assuming that no arbitrage opportunities are present, we may represent the price of a certain asset (  $i$  ) at a given time  $t$  as the sum of weighted payoffs multiplied by their respective probability:

$$P_{i,t} = \sum_{s=1}^k \pi_{t+1}(s) m_{t+1}(s) X_{i,t+1}(s)$$

With  $m_{t+1}(s)$  being the weight for each payoff for the purpose of discounting, and  $X_{i,t+1}(s)$  being the payoff at time  $t+1$  for the asset  $i$ , for a given  $s$  scenario. Considering risky assets (e.g. common stocks) the payoff should include the price movement plus the dividend, both considered at time  $t+1$ , while for zero coupon bonds, the payoff should simply equal the price movement in the asset. On a general basis the returns on the assets are qualified as

$$R_{s,t+1} = \frac{X_{i,t+1}}{P_{i,t}} - 1$$

Considering the safe asset thus, the payoff will be nothing but the above equation considering that the payoff is a certain one, this implying that the payoff only needs to be discounted by the relative weight. This brings our return on the safe asset to be represented as

$$R_{s,t+1} = \frac{1}{\sum_{s=1}^m \pi_{t+1}(s) m_{t+1}(s) X_{i,t+1}(s)} - 1$$

$$R_{s,t+1} = \frac{1}{E_t(m_{t+1})} - 1$$

The return thus depends on the pure discount factor that is applied to the certain payoff. In such a setting, to extract a definition of risk premium we need to compare the safe to the risky asset: peculiarly we may write

$$E_t(m_{t+1}(1 + R_{i,t+1})) = 1$$

Since the gross return, discounted by the stochastic discount factor, is by definition equal to one in expectation. This brings us to consider the covariance between the stochastic discount factor and the gross return, that will be, by definition

$$Cov(m_{t+1}, (1 + R_{i,t+1})) = 1 - E_t(m_{t+1})E_t(1 + R_{i,t+1})$$

The gross return on the risky asset may thus be represented as

$$E_t(1 + R_{i,t+1}) = -\frac{Cov(m_{t+1}, (1 + R_{i,t+1}))}{E_t(m_{t+1})} + (1 + R_{s,t+1})$$

Excess return on this basis will simply be the difference between the safe and the risky asset expected gross return, that is to say

$$E_t(R_{i,t+1} - R_{s,t+1}) = 1 - R_{s,t+1} cov(m_{t+1}, (1 + R_{i,t+1}))$$

Excess return on risky asset with respect to the safe asset will thus be in expectation negatively related to the return on the safe asset and to the covariance between stochastic discount factor and the return on the risky asset. It is now crucial to understand how to interpret this result: if the covariance between the stochastic discount factor and the return of the risky asset is low, then a greater excess return will be expected. This means that assets with high returns when stochastic discount factor is low (i.e. future payoff are diminished by a smaller percentage, and so they are valued more since time is less expensive) will require a higher risk premium. Considering different cases of frequency of observations, forecasting shall be simpler in two points in time which are near, i.e. at higher frequency, when we have the stochastic discount factor which is not varying that much across different points and where the return on the safe asset is near zero, while forecasting will be harder when the frequency is lower and times of observations are further from each other: in this case safe return will be away from zero and stochastic discount factor will be more volatile, requiring higher risk for such uncertainty on a longer horizon. This links directly predictability of risky returns with stochastic discount factors and risk premia, allowing us to say that on longer horizon the assumption of time varying risk premia is much more realistic than the constant risk premia one.

The aim of term structure models is to analyze returns on bonds and see how they move cross sectionally across maturities and along certain periods of time, and it proves useful for example to get some information on expectations about inflation. In performing such analysis it is then comfortable to work with zero coupon bonds, that, having a single cash flow at maturity, allow us to study form a valid theory that can be extended even to ordinary coupon bonds.

The price of zero coupon bonds and their respective yield are usually related by a simple equation as the following

$$P_{t,T} = \frac{1}{(1 + Y_{t,T})^{T-t}}$$

With  $P_{t,T}$  being the price of the zero coupon bond with maturity equal to  $T$  and valued at time  $t$ , and  $Y_{t,T}$  being the respective yield to maturity. Using continuously compounded yields to maturity we can transform the above expression taking logs and get

$$p_{t,T} = - (T - t) y_{t,T}$$

At this point it is crucial to observe that the coefficient through which yields negatively affect price is time to maturity. This means that on a general basis that for a given level of yield, longer maturities will reflect into a lower price, having thus a higher risk premia, mainly due to the higher uncertainty caused by the larger distance between price valuation and payoff.

Passing then to consider holding period returns, a bond with a maturity in  $T$ , indicated as  $r_{t,t+1}^T$  will be represented with

$$\begin{aligned} r_{t,t+1}^T &= p_{t+1,T} - p_{t,T} = - (T - t - 1) y_{t+1,T} + (T - t) y_{t,T} \\ &= y_{t,T} - (T - t - 1) (y_{t+1,T} - y_{t,T}) \\ &= (T - t) y_{t,T} - (T - t - 1) y_{t+1,T} \end{aligned}$$

This is equivalent to say that the difference between yields and returns in a period (of a bond with maturity at  $T$ ), is nothing but the difference between the yield's movement in that period multiplied by a scalar related to maturity. Thus the risk premia will be directly dependent from the respective yields movements, taking also into account time to maturity, and will thus differ from different yields.

At this point we must understand what a no arbitrage condition imposition means. On a general basis no arbitrage requires that there in a given period no profitable riskless investment on an asset, i.e. all assets are appropriately priced in a way that allows no extra market gains without bearing more risk. In our setting this translates into having no excess holding period returns higher than the returns from investing on a bond maturing in the same period taken in consideration. In expectation we will have excess holding period returns from a long maturity bond with respect to a short maturity excess return, where in the latter case excess return will coincide with yield. The excess return will appear as

$$E_t(r_{t,t+1}^T - r_{t,t+1}^1) = E_t(r_{t,t+1}^T - y_{t,t+1}) = \varphi_{t,t+1}^T$$

$$E_t(r_{t,t+1}^T) = y_{t,t+1} + \varphi_{t,t+1}^T$$

From this equation, reintroducing the log price of the bond related to the returns, using the properties of the logarithm and the fact that price at time  $t$  equals the price in the following period discounted by the rate of return, the log price of the bond evaluated at  $t$  with maturity at  $T$  can be represented as

$$p_{t,T} = p_{t+1,T} - r_{t,t+1}^T$$

Hence the yield of the bond will appear as

$$y_{t,T} = \frac{1}{(T-t)} \sum_{i=0}^{n-1} E_t(r_{t+i,t+i+1}^T)$$

$$= \frac{1}{(T-t)} \sum_{i=0}^{n-1} E_t(y_{t+i,t+i+1} + \varphi_{t+i,t+i+1}^T)$$

In the end, given the no arbitrage condition, the yield of a given bond with a fixed maturity will appear as the mean of the future expected one period rates and the respective risk premia.

### 2.2.1

### FORWARD RATES

An important concept to conceive to well understand term structure theories is the one of forward rates. As a general definition, we may say that forward rates are the yield that an agent would agree to get today for an investment at some future time  $t$  over the period following  $t$

until maturity  $T$ . these kind of investments can be summarized by the yield curve, as a ten year zero coupon bond may be decomposed into a one year zero coupon bond and other nine forward rates. The strategy for the replication of the forward rate is usually defined as the one who sells at time  $t$  a bond maturing at time  $t'$ , and buys a bond at time  $t$  maturing at time  $T$ . Through this reinterpretation given by the investment strategy, we are able to write down the forward formula, that is

$$f_{t,t',T} = \frac{(T-t)y_{t,T} - (t'-t)y_{t,t'}}{T-t'}$$

Where at the numerator we have the log price of the same forward, and maturity at denominator. This means that for example, a forward rate for a bond investing at time  $t+1$  and maturing at  $t+2$  equals two times the yield from time  $t$  to time  $t+2$  minus the one period yield. Using such a rule we may recursively describe any maturity yield using forward rates and the one period yield, thus we could write

$$y_{t,t+n} = \frac{1}{n}(y_{t,t+1} + f_{t,t+1,t+2} + f_{t,t+2,t+3} + \dots + f_{t,t+n-1,t+n})$$

And recalling that

$$y_{t,t+n} = \frac{1}{n} \sum_{i=0}^{n-1} E_t(y_{t+i,t+i+1} + \varphi_{t+i,t+i+1}^T)$$

We come to the conclusion that

$$f_{t,t+i,t+i+1} = E_t(y_{t+i,t+i+1} + \varphi_{t+i,t+i+1}^T)$$

This last consideration equates future short rate to forward rate and explicitly puts forward that forward rates are nothing but the expected short yield plus a given risk premium.





# 3

## PRICING BONDS IN CONTINUOUS TIME

---

Pricing zero-coupon bonds within the economy through a model is what is usually required to a term structure model that wants to create, or somehow emulate, the relationship between bond of different maturities and between yields and maturities. Zero-coupon bonds are obligations paying a given capital, usually named *principal*, after a given period, but giving no intermediate payments (*coupons*) during the same period. To model such products and their yields usually the principal is taken to be equal to one, to have a common ground for comparison.

Of course to keep comparability among bonds, time to maturity will define bond price. Even contract with different maturities but with equal time to maturity at a given date will be considered the same. We assume each bond is traded on the market at a given price  $P^\tau$ , that defines the yields of each bond with a given maturity in this way:

$$hpr_{t,t+n}^\tau = \log P_{t+n}^{\tau-n} - \log P_t^\tau$$
$$y^\tau = \frac{hpr_{t,t+\tau}^\tau}{\tau} = - \frac{\log P_t^\tau}{\tau}$$

That is, independently from the specific term of the contract, the relevant features of the instruments we are considering are no-coupon and common time to maturity: note that if it was otherwise arbitrage opportunities could easily been exploited. The specification of the yield given above tells nothing but that the yield is a particular case of holding period return, when the bond is kept until maturity.

The short rate, that in ATSMs plays a central role for specification, is theoretically the limit of the yield as time to maturity shrinks and goes towards zero. This means that the short rate is resembling an hypothetical *instantaneous rate*. Another key definition is excess holding period returns, that is the holding period return in excess of risk free return, defined as the difference between the *hpr* of a bond and the *hpr* of another bond that is maturing at the end of the holding period.

The method usually employed to price bonds uses a corresponding risk neutral probability measure, most of the times identified with  $Q^*$ , and it is in this probability space that risk neutral pricing is applied. Risk neutral measure is an artificial one, where prices in the economy are

payoffs discounted at the riskless rate. The key for the difference between the data-generating measure (where we observe the events) and the artificial risk neutral measure is the computation of probability weights. Such probability measure, mostly used for derivative pricing, assures that arbitrage is not possible: under  $Q^*$  prices are martingales, that is, the expectation of prices under such measure, conditional on past history of the same asset, is equal to the last observation for the price of the asset, a well defined model in this sense is Duffie (2001). Thus generally what is used as a general formula to price bonds in continuous time is

$$P_t^\tau = E_t^* [\exp(-\int_t^{t+\tau} r_u du)]$$

This obviously means that the price at time  $t$  of a bond with maturity  $\tau$  is depending on the expectation under the corresponding martingale (i.e. risk neutral measure) directly on the movement of the short term rate from the time of pricing to the maturity of the bond, this is equivalent to say that long term rates are nothing but the mean of expected future short term rates. This seems quite interesting since we can easily understand that under the martingale measure excess holding period returns are equal to zero. Piazzesi (2006) well defines this passage by allowing some error in notation and putting it as

$$E_t^* \left[ \frac{dP_t^\tau}{P_t^\tau} \right] = r_t dt$$

That is to say that under the risk neutral measure the instantaneous movement of the price of the bond is perfectly identified by the movement of the short rate. The important thing thus becomes the passage from data-generating measure to the risk neutral probability measure, this is what determines the fundamental in the pricing of the bond, and this is precisely determined by the so called *term premia* definition.

Generally term structure model in continuous time setting thus include two main stages:

- 1) The change of probability measure from the data generating measure  $Q$  to the risk neutral one,  $Q^*$ ;
- 2) A specific description of the movement of the short rate  $r$  under the new martingale measure  $Q^*$ .

One famous class of models, named *factor models*, by the way substitutes point 2) with a different assumption. The dynamics of the short rate  $r$  is described through a relationship with a factor (as in Vasicek) or with a vector of factors, usually named *state variables*, and is usually

set to be a Markov continuous process. A Markov process is defined as a process in which expectation conditional on the past is equal to the last observation, it is a memoryless process. Hence the conditional expectation at time  $t$  under the martingale measure we showed above becomes a generic function of time to maturity and of the state vector. For each time to maturity of a bond we will have a price defined by this function and by the relationship with the state variables that are included in the model.

$$P_t^\tau = F(x_t, \tau)$$

Furthermore a dependence on time may be included, bringing the relationship to be time varying for different factors

$$P_t^\tau = F(x_t, t, \tau)$$

This depends on how term premia are defined within the model. Some research ignored term premia, some other considered them as time constant, and other took them as time varying: there are different models with different setups and term premia specification is usually set independently.

Even if it will not be our case, it is important to quote how usually ATSMs price bonds in the economy. In continuous time the key tool that is used is Ito's Lemma, that, for a given Ito process and a given smooth function of such Ito process, even nonlinear, provides a dynamic for the price and turns the problem into a PDE problem usually solved through the Feynman-Kac approach. The final formula for bond price thus, in most ATSMs, is a PDE derived from the relation between yields and state variables and risk premia specification.

### 3.1

#### THE LOCAL EXPECTATION HYPOTHESIS

---

The local expectation hypothesis, frequently indicated as LEH, is one of the basic assumption that, mostly in past, were stated as a starting point in term structure models. Specifically, the LEH allows the equation

$$P_t^\tau = E_t^* [\exp(-\int_t^{t+\tau} r_u du)]$$

to hold under the data-generating measure. This means that it is assumed that in the probability measure where observations are made, yields are the log expectation of future short rates until maturity, and we are hence allowed to write down

$$y_t^\tau = -\log E_t[\exp(-\int_t^{t+\tau} r_u du)]/\tau$$

That corresponds to the log price divided by the number of periods, given that the  $Q$  and the  $Q^*$  probability measures are the same. On a general basis thus, excess returns on longer maturity bonds are zero, since every yield is exactly the expected payoff for that maturity even if we are not in an artificial risk neutral measure. This differs from the Expectation Hypothesis (EH), that describes yields as the expectation of average future short rates

$$y_t^\tau = E_t[S]/\tau$$

Assuming normality of the short rate we can see that the local expectation hypothesis gives the yield as the logarithm of a lognormal distribution, this means that LEH can be rewritten as

$$y_t^\tau = \frac{E_t[S]}{\tau} - \frac{1}{2} \text{var}_t[S]/\tau$$

Hence the two assumptions differ for the variance term. On the empirical ground different papers were written on the compatibility of such hypotheses with no-arbitrage assumption, as Cox Ingersoll Ross<sup>1</sup> (1981) and Fisher (1998)<sup>2</sup>. Moreover for different specification of the short rate, EH and LEH provide different shapes of yield curve, since each of these hypotheses gives a different relationship between yield and maturity. In the next subsection we will briefly focus on the bonds price PDE in the case in which LEH is present or not.

### 3.1.1

#### PDE FOR BOND PRICES: THE ROLE OF LEH

As we mentioned above, in ATSMs usually the change of measure from  $Q$  to  $Q^*$  has to be provided, but if LEH is assumed then no change will be needed since the two measures will coincide. Even if LEH is a pretty strong hypothesis, that has been denied as counterfactual, it does have some advantages. Working directly in the data generating measure in fact can allow

---

<sup>1</sup> J.C. Cox, J.E. Ingersoll, S.A. Ross, 1981, "The relation between forward prices and future prices", journal of Financial Economics 9, 321-346

<sup>2</sup> Fisher, M., 1998. A Simple Model of the Failure of the Expectations Hypothesis, Working paper (Federal Reserve Bank of Atlanta, Atlanta, GE)

some insight on parameters in the same probability space of the observations, and needs no complex changes of measure within the model.

To have a brief idea of what this means we will consider a Markov process  $x$  solving the SDE

$$dx_t = \mu_x(x_t)dt + \sigma_x(x_t)dz_t$$

with drift  $\mu_x$ , volatility  $\sigma_x$ , and  $z$  being a  $Q$ -brownian motion. Among such processes the so called Gaussian processes have affine drift and constant volatility, but in general the drift and the volatility are time –homogeneous. The key idea is that a partial differential equation is built starting from the basic concept that the expression for price

$$P_t^\tau = E_t^* [\exp(-\int_t^{t+\tau} r_u du)]$$

is the solution for the bond price, that depends on state vector and maturity, i.e. the bond price is  $F(x, \tau)$ . From the pricing equation above we know that the price at time  $t$  for a given maturity is equal to the payoff, this means that when time to maturity runs out, and thus  $\tau = 0$ , we will have  $F(x, 0) = 1$ . Again, from the pricing equation we know that price is strictly positive (exponential function), and so will be even  $F(x, \tau)$ , that will also be an Ito process with some drift  $\mu_F(x_t, \tau)$  and some volatility  $\sigma_F(x_t, \tau)$ . As a final step we will have that the initial Local Expectation Hypothesis will imply that the drift of the price, function of state variables and maturity, will be nothing but the short rate  $r$ . On a general basis thus starting from a pricing equation that gives the price as an expected payoff, being the price a function of state variables and imposing LEH, we come to a model that can state a relationship between prices of bonds and state variables. Specific resolution of PDE may bring some trouble on an empirical ground, but for what concerns our paper this is not strictly needed.

A different case, and with no doubt a more realistic one, is the one in which Local Expectation Hypothesis is not assumed. This means that the data-generating measure is not taken to coincide with the risk neutral measure, hence a new probability space will have to be define and a “translation” mechanism between the two spaces will be needed to compute prices. If with LEH the drift of the price function was equal to the short rate, i.e.

$$\mu_F(x, \tau) = R(x)$$

now the drift will be equal to the short rate only under the martingale measure  $Q^*$ , i.e.

$$\mu_F^*(x, \tau) = R(x)$$

and a passage from the two measures will be needed. The simplest case for the dynamics of  $x$  will be taking it as a single state variable, as the first ATSMs did (e.g. Vasicek, 1977)<sup>3</sup>: in such models the state variable considered was simply the present short rate  $r = x$ . In such case the distribution of the short rate would be shifted, consequently form an adjustment for risk. The mean of the state variable will hence change passing through the measures, while volatility would stay the same, i.e.

$$\sigma_x^* = \sigma_x$$

The whole process for the construction of PDE will now have to be derived in the risk-neutral measure and only after will be translated back to the data generating measure.

This change of measure is usually performed involving a density  $\xi$ , a strictly positive martingale starting from  $\xi_0 = 1$  and moving according to the Stochastic Differential Equation

$$\frac{d\xi_t}{\xi_t} = -\sigma_\xi(x_t)dz_t$$

This density is assured to be a martingale since it solves Novikov condition, i.e. the expectation for the solution of  $\xi_t$  is finite. Once this is satisfied, the Q-distribution can be twisted into the risk neutral probability measure through the Girsanov theorem, that bring the dynamics of  $x$  to

$$dx_t = (\mu_t^*(x_t) - \sigma_t^*(x_t)\sigma_\xi^*(x_t)^T)dt + \sigma_t^*(x_t)dz_t$$

Where we can note that  $x$  keeps the same volatility under both measure: this is known as the *diffusion invariance principle*.

---

<sup>3</sup> Vasicek, O., 1977. An Equilibrium Characterization of the Term Structure, Journal of Financial Economics, 5, 177-88.

## 3.2

## AFFINE TERM STRUCTURE MODELS

---

The second stage of term structure model quoted in the above sub-chapter is the main concern for what defines affine term structure models. In fact generally ATSMs have as a common and characterizing feature the functional form that relates the short rate and the state variables, and the type of diffusion that these state variables show. The functional form intended is of course affine form, which is characterized by a constant plus a linear term, and this is applied:

- 1) To  $R(x)$

Where we will have a short rate as function of state variables in an affine form like

$$R(x) = a + b^T x$$

- 2) To the diffusion of  $x$  under the risk neutral measure, where both the drift  $\mu_x^*(x_t)$  and the variance matrix  $\sigma_t^*(x_t)\sigma_t^*(x_t)^T$  will be affine functions.

The first requirement appears pretty simple, being  $a$  a real constant and  $b$  some vector of coefficients, with length equal to the number of factors considered within the model. ATSMs with one state variable will thus have  $a$  equal to zero and  $b$  coefficients vector of the type  $[1 ; 0_{N-1}]$ .

The second requirement instead needs a more detailed explanation.

### 3.2.1

### AFFINE DIFFUSIONS

To explain affine diffusions it is helpful for us to start from a general case, that may result simpler, that provides the imposition of Local Expectation Hypothesis, this correspond again to assume that risk neutral pricing can be done in the data-generating probability measure, since the assumption on price as expected exponential short rate holds in such space.

Assuming that the  $x$  vector has an affine diffusion means to assume that the dynamics of such process is

$$dx_t = \mu_x(x_t)dt + \sigma_x(x_t)dz_t$$

Where the coefficients for drift and volatility are affine functions of  $x$ , and specifically

$$\mu_x(x_t) = k(\bar{x} - x)$$

$$\sigma_x(x_t) = \Sigma s(x)$$

Where  $s(x)$  is an  $N \times N$  matrix with  $\sqrt{s_{0i} + s_{0i}^T x}$  for each  $i$ -th diagonal element, while  $\Sigma$  and  $k$  are matrices of constants.

To better understand what this means it may be of some use to make clear some consequences that in the univariate case are of simple intuition.

First of all, the drift of the affine diffusion is nothing but a mean reverting process. In fact the error observed at each time, intended as the distance of the state variable from its own mean, is corrected with speed of adjustment  $k$  (taking it between 0 and 1): if the level of  $x$  is above the mean, the movement  $dx$  brings  $x$  towards its mean by a fraction of the error equal to  $\frac{1}{k}$ . In the meanwhile, this movement of mean reversion will be disturbed by the stochastic part of the infinitesimal movement  $dx$ . In fact to the mean reversion component a random shock  $dz_t$ , distributed as a normal with mean 0 and variance  $dt$ , will add some noise via the coefficient  $\sigma_x(x_t)$ . Thus depending on whether the volatility is allowed to be time varying or not, the random normal shocks will translate more or less during periods of high or low volatility into the infinitesimal movement  $dx$ .

The most common types of affine processes are Gaussian and Square Root processes, which are characterized by the assumptions on the stochastic coefficient (volatility matrix). The first type of process has constant variance matrix: this means that  $\sigma_x(x)\sigma_x(x)^T$  is time homogeneous, and thus the disturbance term affects the movement of the state variables in a stable way throughout time. On a mathematical basis this translate into  $s(x)$  being an identity matrix, since  $\Sigma$  is a free parameters matrix. We will thus have a process described by the stochastic differential equation

$$dx_t = k(\bar{x} - x)dt + \Sigma dz_t$$



Square root processes instead allow a linkage between the variance matrix and the state variables. This brings along heteroscedasticity and thus needs for some requirements to avoid that zero is not reached by the process.

The process will appear as

$$dx_t = k(\bar{x} - x)dt + \Sigma\sqrt{x_t}dz_t$$

Variance will thus be  $\Sigma^2 x_t$ , that may be not positive definite. If the process  $x_t$  hits zero in fact, variance will do as well: to avoid this the Feller condition  $k\bar{x} > \frac{1}{2}\Sigma^2$  ensures that the drift part is high enough to make the process go away from zero. Moreover, an important feature to note is that the variance of the process in this case will be always proportional to the value of the process at time  $t$ .

At this point it is really important to note that Gaussian processes directly depend on their past history. Peculiarly this feature makes the infinitesimal movement of the process  $x$  to depend on a deterministic part that pushes towards the mean value, and a stochastic part that, even with constant volatility coefficient, disturbs this movement. In a discrete time setting this translate in a Vector Autoregressive stochastic process, that makes each value at each time  $t$  to depend directly on its past. This will be fundamental for us since the model we will implement will be characterized by Gaussian processes in a discrete time setting as well.



# 4

## INTEREST RATES TERM STRUCTURE: THE MODEL

---

### 4.1

#### THE MACRO - FINANCE MODEL

---

In this chapter we are going to describe the general approach that has been used in the paper. First of all the base model we are taking as main inspiration is the one by Ang Piazzesi (2003)<sup>4</sup>, where the macroeconomic models and the finance models were mixed to get a no arbitrage autoregression of the yield curve, introducing macro factors in an affine term structure model. The next session is describing the rational, the main mathematical results of Ang Piazzesi (2003) and finally the estimation method used.

The main motivation for the model by Ang and Piazzesi is the fact that the macro literature on the yield curve only describes the link between the yield considered and the macroeconomic variables. On the other side finance models, and in particular affine term structure models, working most of times in a continuous time setting and studying the deep meaning of the existing difference between the different maturities yields, were able to forecast all the possible maturities. By the way the link between yield curve and macroeconomic variables was left on the table by these models, or at least was not studied explicitly. On the empirical side this is done by using a Gaussian ATSM and adding macroeconomic variables to it.

The macro variables included in the model, to ensure estimation, are reduced to two measures: the first one is a proxy for inflation, the other one is a proxy for real activity. The construction of such variables is made by way of the principal component analysis, taking seven different

---

<sup>4</sup> Ang, Andrew and Monika Piazzesi (2003) "A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables." *Journal of Monetary Economics* 50,745–787.

measures of activity and real activity, dividing them into two groups and taking the first component of the principal component analysis.

This kind of analysis, starting from the variance-covariance matrix of original variables and the relative eigenvalues, produces a series of linear combinations of the original series that give as output a new series of variables called principal components, where the variation of original variables is transferred in different percentages to the new ones. Taking the first principal components thus leads to have a kind of synthetic variable explaining most of the times more than 90% of the variation of initial variables. In the case of Ang and Piazzesi the initial variables that are taken are consumer and producer price index, spot market commodity prices, unemployment rate, the index of Help Wanted Advertising in Newspapers, the growth rate of employment and the growth rate of industrial production.

The interesting approach used starts from a Taylor Rule, where the short rate dynamics is dependent on Gross Domestic Product and Inflation Rate<sup>5</sup> measures. The unobservable variables then are nothing but the error of such Taylor Rule, that is they are independent from macroeconomic variables effects. Thus while a simple Taylor Rule appears as

$$r_t = a_0 + a_1 f_t^o + v_t$$

Ang Piazzesi (2003) will describe the short rate as

$$r_t = \delta_0 + \delta_{11} X_t^o + \delta_{12} X_t^u$$

This kind of description of the short rate dynamic is one of the base equations of the affine term structure model, together with the measurement equation, the description of the Radon Nikodym derivative and the vector autoregression of the model factors.

The state dynamics on the other hand is described by the following autoregression

$$X_t = \mu + \Phi X_{t-1} + \Sigma e_t$$

---

<sup>5</sup> Actually, the measures used are GDP and Inflation gaps, that is the difference between the two macro variables and the relative target measures.

Where  $X_t$  is the vector of factors describing the yield curve.

Moreover a pricing kernel is performed to assess the market prices of risk. This kind of assumption made at the base at the model allows to assume the no-arbitrage and thus guarantee the existence of an equivalent martingale measure. This means that every asset in the economy perfectly compensates the risk beared, that is inherent in the asset. This kind of statement is used to assure that no extra gain is made in the economy beyond risk free rate and risk premium.

Applying this concept to the yield curve means to state that the excess returns on the risk free rate, that is the excess returns on the one month zero coupon bond, are determined by the risk that the agent in the economy is bearing. The whole model is dependent upon such rational: why are longer maturities always showing higher returns? Because they are reflecting the risk caused by the enhanced uncertainty, due to the further distance of maturity. This uncertainty will depend on some variables, that in our case are included in vector  $X_t$ . In fact starting from the autoregression of  $X_t$ , a vector  $e_t$  of errors will be left out, and it will represent the unexplainable portion of the state variables taken in consideration. This is in fact what is inserted in the Radon Nykodim derivative, being the derivative that describes the link between the observation probability measure and the risk neutral probability measure:

$$\xi_{t+1} = \xi_t \exp\left(-\frac{1}{2}\lambda_t' \lambda_t - \lambda_t' e_{t+1}\right)$$

This means that the difference between data generating probability measure and the risk neutral probability measure, is dependent on the unexplained portion in the autoregression of the state vector by way of some vector  $\lambda_t$ , that creates the recursive series  $\xi_t$ . The vector  $\lambda_t$  can be thus parametrized by an affine transformation as

$$\lambda_t = \lambda_0 + \lambda_1 X_t$$

Where  $\lambda_0$  and  $\lambda_1$  will be the market prices of risk. In fact the diverse  $\lambda_t$  will be the multipliers from uncertainty in economy  $e_t$ , that will bring to the Radon Nykodim derivative. But being this derivative recursive, the first two  $\lambda_t$  will be the most important and they can be interpreted as the prices that market attributes to uncertainty, i.e. risk.

The pricing kernel thus is completely defined as

$$m_{t+1} = \exp(-r_t) \xi_{t+1} / \xi_t$$

And hence, substituting  $r_t$  with the short rate dynamics we come to

$$m_{t+1} = \exp\left(-\frac{1}{2}\lambda_t'\lambda_t - \delta_0 + \delta_1 X_t - \lambda_t' e_{t+1}\right)$$

And we get to the description of the pricing kernel as a function of the past values of the state variables and the transformation of the errors describing uncertainty.

To apply this to the bond prices formula we start from the consideration that uses the definition of pricing kernel

$$E_t(m_{t+1}R_{t+1}) = 1$$

That is the one period rate (risk free rate) multiplied by the factor equal to the pricing kernel, gives unity. In this way prices of bonds can be computed recursively by using

$$p_{t+1}^{n+1} = E_t(m_{t+1}p_{t+1}^n)$$

Since the pricing kernel shapes the behaviour of an additional period of maturity return. Moreover, within the context of the Ang Piazzesi (2003) model, the prices of bonds are given by an affine transformation, named measurement equation, of the state vector that looks like

$$p_{t+1}^n = \exp(\bar{A}_n + \bar{B}_n' X_t)$$

For the price at time  $t+1$  of a zero coupon bond maturing in  $n$  periods. Where the two coefficients can be found algebraically<sup>6</sup> and result equal to

$$\bar{A}_{n+1} = \bar{A}_n + \bar{B}_n'(\mu - \Sigma'\lambda_0) + \frac{1}{2}\bar{B}_n'\Sigma\Sigma'\bar{B}_n - \delta_0$$

$$\bar{B}_{n+1}' = \bar{B}_n'(\Phi - \Sigma'\lambda_1) - \delta_1'$$

Starting from

$$\bar{A}_1 = -\delta_0$$

$$\bar{B}_1' = -\delta_1'$$

---

<sup>6</sup> For algebraic derivation of the coefficients please refer to Ang Piazzesi (2003) Appendix.

That is nothing but the dynamics of the short rate. Looking at the two equation for the coefficients of the measurement equation we can see that at each time  $t$  a quantity depending on autoregressive parameters and on market prices of risk is added and in particular the short rate dynamics mean and slope are taken away at each added period.

Yields values can be finally computed since

$$y_t^n = -\frac{\log p_t^n}{n}$$

$$y_t^n = A_n + B_n' X_t$$

With  $A_n$  and  $B_n$  being equal respectively to  $-\bar{A}_1/n$  and  $-\bar{B}_1/n$ .

Given this general theoretical framework, already used in many other works<sup>7</sup> needs some more specifications in this model. In fact, beyond being an Affine Term Structure Model, this model includes both observables and unobservables in the state vector  $X_t$ , being a kind of Taylor Rule with specified errors, equal to the unobservables factors: the stronger assumption made is the one of independence of macro and latent factors.

This of course goes hand in hand with the Taylor Rule interpretation of the model, but on the empirical basis, it is assumed that no link is present between macro and the latent factors of the term structure. This would go against the evidences of predictive power of macro movements by the yield curve and of course of the link between short rate monetary policy and macroeconomic variables<sup>8</sup>. That could be seen as the main drawback of the model, despite the forecast from it proves to be rather effective.

Another interesting feature of the model are the Risk Premia assumptions. In such a setting  $\lambda_t = 0$  would mean assuming that data generating measure and risk neutral measures coincide: that is exactly the Local Expectation Hypothesis case, explained in the previous sections. In such a case risk premia would be null, since the coefficients would collapse to have constant coefficients in the measurement equation, and thus the mean returns per month of holding a one year zero coupon bond or a ten year coupon bond would be the same. In some models, like macro models (where simple VAR on macro variables is performed) this is assumed, in others

---

<sup>7</sup> A rather complete classification of such models can be found in Dai, Q., Singleton, K., 2000. Specification analysis of affine term structure models. *Journal of Finance* 55, 1943–1978.

<sup>8</sup> Such evidences can be found for example in Harvey, C.R., 1988. The real term structure and consumption growth. *Journal of Financial Economics* 22, 305–333.

the risk premia are assumed to be constant (and thus all equal to  $\lambda_0$ )<sup>9</sup>. This seems to be thus a crucial point, since it would significantly change the final yield curve: the no arbitrage assumption and the risk premia specification do characterize differences among maturities and so the whole shape of the curve.

On the estimation side what is performed is a change from a system of yields and observables into a system of observables and unobservables. Studying the density functions of yields and macro variables, unobservables can be inferred. The main point is to start from the relationships between yields and observables, to describe some new variables basing on a restricted number of yields (three in Ang Piazzesi, 2003), and analyse the links between the newly found unobservables, together with macro, and the yields that are left: this is referred to as the Chen and Scott (1993) method<sup>10</sup>. To do this a two steps approach is used:

- 1) First of all, Ordinary Least Squared (OLS) is used to find the main coefficients  $\delta_0$  and  $\delta_{11}$  of the short rate dynamics and the autoregressive coefficients  $\mu$  and  $\Phi$ . This of course are vectors of parameters, that contain parameters corresponding to unobservable factors. The point is that unobservable factors are found simultaneously in the model, thus in this first step only the parameters corresponding to observables (macro factors) are computed and kept fixed in the subsequent estimation step
- 2) Holding the previously found parameters fixed, a maximum likelihood function is estimated to assess values for the remaining parameters (namely market prices of risk and the parameters corresponding to unobservable factors).

On a more practical base, the following equations are used

$$r_t = \delta_0 + \delta_{11}X_t^o + \delta_{12}X_t^u$$

$$X_t = \mu + \Phi X_{t-1} + \Sigma e_t$$

$$\bar{A}_{n+1} = \bar{A}_n + \bar{B}_n'(\mu - \Sigma'\lambda_0) + \frac{1}{2}\bar{B}_n'\Sigma\Sigma'\bar{B}_n - \delta_0$$

$$\bar{B}_{n+1}' = \bar{B}_n'(\Phi - \Sigma'\lambda_1) - \delta'_1$$

$$\lambda_t = \lambda_0 + \lambda_1 X_t$$

$$y_t^n = A_n + B_n'X_t$$

<sup>9</sup> See for example Vasicek (1977)

<sup>10</sup> Chen, R.R., Scott, L., 1993. Maximum likelihood estimation for a multi-factor equilibrium model of the term structure of interest rates. *Journal of Fixed Income* 3, 14–31.



Exactly in this order estimation is provided. Starting from the short rate dynamics and the vector autoregression, initial values for macro factors are estimated. From these vectors we obtain some initial values for the vectors of  $A_n$  and  $B'_n$ , and  $X_t^u$  time series are computed: following this first stage the likelihood function will begin to run, estimating again at each iterative round new vectors of parameters  $A_n$  and  $B'_n$  also considering the new unobservables  $X_t^u$ , until the likelihood function, based on the probability density functions of observables, unobservables and residuals is optimized for a given vector of parameters  $\theta$ , including the autoregression parameters  $(\mu, \theta)$ , short term dynamics parameters  $(\delta_{12})$  corresponding to unobservable factors, and market prices of risk  $(\lambda_0, \lambda_1)$ . This whole system will finally provide a given vector of optimizing parameters, that will return some definite measurement equations, one for each maturity we may need to know. The interesting aspect of such model is the fact that studying autoregressive portions of factors, permits to isolate the effects that the factors at each time  $t$  have on their future values, and the effect that will affect the short rate. Then there will of course be an additional unobservable effect determining the short rate, represented by the unobservable factors and their relative parameters. Moreover longer maturities will suffer some additional effect, due to higher risk, and thus mainly determined through the estimation of market prices of risk, that return the additional premium from each unit of risk, that in this case is an additional month of maturity.

In such a way a given prediction for short rate is given based on the vector of factors, and the entire yield curve for potentially every possible maturity (on an empirical basis discrete maturities with one month differences) can be predicted. Of course the entire model is based on the effective representation of short rate through adequate choice for the state vector and the specification of risk premia. In the case of Local Expectation Hypothesis would disregard the second additional effect for longer maturities, and thus would set market prices of risk to zero. The LEH in fact would assume that for example, a twelve months zero coupon bond could be replicated through a rolling strategy buying twelve subsequent one month zero coupon bonds, thus holding a single twelve months bond would give no excess return, that is market prices of risk would be zero. Below we report the likelihood function provided directly by the Appendix of the original paper by Ang and Piazzesi (2003): the four elements composing such likelihood function are the density functions of the variables that linearly combined through the parameters of measurement equation return the values for yields.

$$\begin{aligned}
\mathfrak{L}(\theta) &= \prod_{t=2}^T f(Y_t, X_t^o | Y_{t-1}, X_{t-1}^o) \\
\log(\mathfrak{L}(\theta)) &= \sum_{t=2}^T -\log |\det(J)| + \log f_x(X_t^o, X_t^u | X_{t-1}^o, X_{t-1}^u) + \log f_{u^m}(u_t^m) \\
&= -(T-1) \log |\det(J)| - (T-1) \frac{1}{2} \log(\det(\Sigma \Sigma')) \\
&\quad - \frac{1}{2} \sum_{t=2}^T (X_t - \mu - \Phi X_{t-1})' (\Sigma \Sigma')^{-1} (X_t - \mu - \Phi X_{t-1}) \\
&\quad - \frac{T-1}{2} \log \sum_{i=1}^{N-K_2} \sigma_i^2 - \frac{1}{2} \sum_{t=2}^T \sum_{i=1}^{N-K_2} \frac{(u_{t,i}^m)^2}{\sigma_i^2}
\end{aligned}$$

The aim of maximization will be, conditionally on values for the state vector  $X_t$  (containing both observable macro factors and unobservable latent factors), to find the vector of parameters  $\theta$  that has the highest probability to return the correspondent vector of yields  $Y_t$ .

Once this kind of general framework is clear we can better define the Chen and Scott (1993) method. The fundamental concept is the division of the yield into two groups, one of which will in a first step model define the unobservable factors, extracting them from the general Taylor Rule model. The unobservable factors are thus nothing but the residual unexplained portion of the yields chosen for such extraction: taking the three shortest maturities means that the variables taken as unobservables will completely define the shortest yields and that the market prices of risk will be defined, beyond macro factors, by the portion of such yields that is not explained by the values of the same yields and the macro factors at  $t-1$ . This is a pretty complex argument but in effect the definition of market prices of risk is made on an affine function of the state vector, and if the state vector includes variables extracted from the yields this means that we are netting the measure of market price of risk from the autoregressive effect and looking at the unexplained portion. Here is why we may propose, a further analysis on such point: a hint for the chosen yields may be to pick one short, one medium and one long rate for unobservable factors definition, so that we may have the autoregressive effects on three completely diverse maturities within the model.

The general results of the Ang and Piazzesi (2003) can be summed up by a substantial contribution to forecasting by macroeconomics variables, peculiarly on an out of sample basis. Different models for short rates dynamics are used since many times too complex models are outperformed by parsimonious ones. Starting from a random walk and going through an unconstrained and a constrained VAR, that is constrained considering no-arbitrage

implications, the final findings show that a model with macro factors, not including lagged values for macro factors, outperforms the others.



# 5

## THE MODEL

### 5.1

#### YIELDS DATA

The data we will take in consideration in our empirical model will be US interest rates at fixed selected maturities, available on the St. Louis Federal Reserve website<sup>11</sup>. The maturities available here are 1 month, 3 months, 6 months, 1 year, 2 years, 3 years, 5 years, 7 years, 10 years. As declared on the FED website<sup>12</sup> yields on treasury nominal securities at constant maturity are interpolated by the U.S. Treasury from the daily yield curve for non-inflation-indexed Treasury securities. The curve is based on the closing market bid yields on actively traded Treasury securities in the over-the-counter market. These market yields are calculated from composites of quotations obtained by the Federal Reserve Bank of New York. This method provides a yield for a 10-year maturity, for example, even if no outstanding security has exactly 10 years remaining to maturity. What is reported in the graph below are the selected nominal yields calculated through the cited interpolation and taken at a monthly frequency.

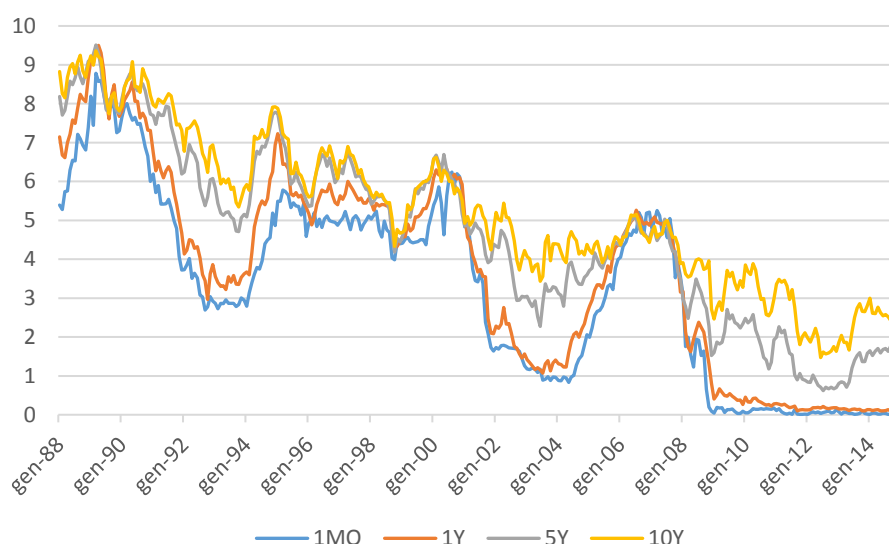
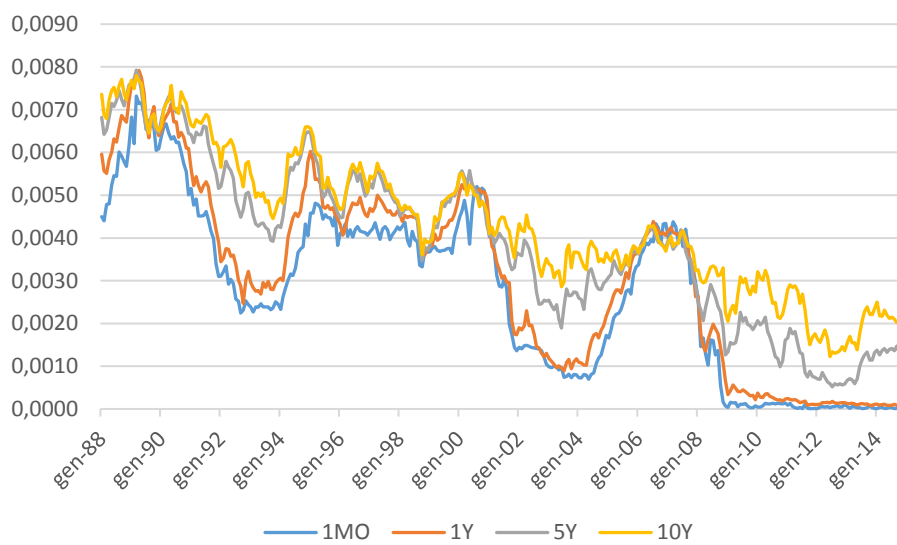


Figure 1 Representative Nominal Yields from the Database

<sup>11</sup> [www.fred.stlouisfed.org](http://www.fred.stlouisfed.org)

<sup>12</sup> [www.federalreserve.gov/releases/h15/current/h15](http://www.federalreserve.gov/releases/h15/current/h15)

As we can easily see from the above graph the yields had a substantial downward movement from the end of eighties until 2014. This period came after the Alan Greenspan took the office at Federal Reserve Bank and brought as a main action in monetary policy the lowering of short rates as a response to all the crisis that United States faced, going from the Russian crisis, to the 2001 Dot-Com bubble burst until 2008 recession. The dataset we will be using considers this periods, with particular focus on the 2008 crisis, that was enhanced by the credit expansions that allowed subprime mortgages and CDO financial market takeover until Lehman Brothers bankruptcy.



*Figure 2 Annualised Representative Yields*

Since we will be dealing with annualised changes to macro variables, the yields are annualised and thus scaled down by a substantial percentage.

## 5.2

### MODEL 1: A SINGLE FACTOR NO ARBITRAGE MODEL

As a first model we will take a single factor model. Such model can be seen as similar to a Vasicek (1977) model, where the state factor used is simply the short rate: in our case the proxy used will be the one month rate. Using data from January 1988 to December 2014, we will give in this model a general example of how this kind of affine term structure models works, we will produce some in-sample forecasts of the model going from January to December 2014. These forecasts for the entire yield curve will be the measure for comparison with next models.

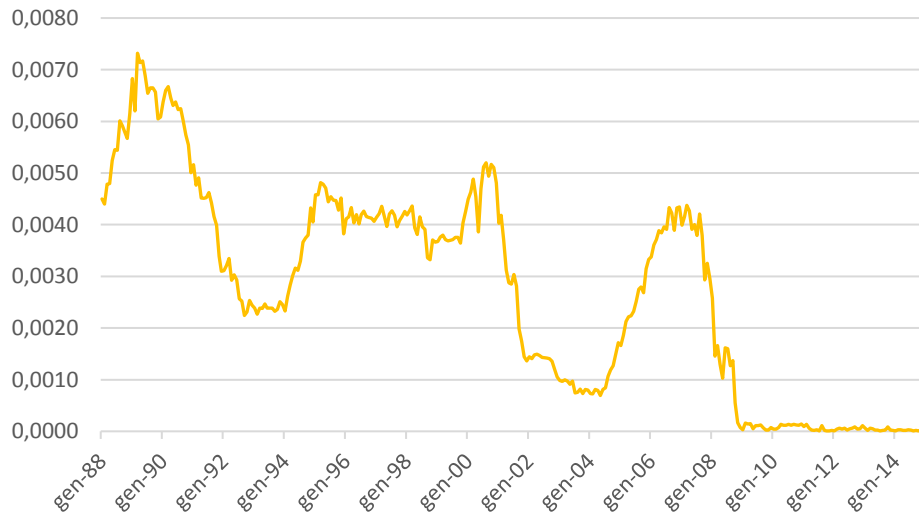


Figure 3 The 1-Month Zero Coupon Bond Yield

The one month yield for zero coupon bonds above graphed will thus constitute our vector  $X_t$  in the equations of the model. Starting from the base vector 12 lag autoregression

$$X_t = \mu + \Phi X_{t-1} + \Sigma e_t$$

We will recover parameters in the  $\mu$ ,  $\Phi$  and  $\Sigma$  through Ordinary Least Squares method. The dynamics of the short rate instead in this first model collapses to an identity since the empirical proxy for the instantaneous rate in our analysis is actually the one month yield, Hence in the equation

$$r_t = \delta_0 + \delta_{11}X_t^o + \delta_{12}X_t^u$$

The parameters  $\delta_0$  and  $\delta_{12}$  will go down to zero, while the parameter  $\delta_{11}$  will be simply equal to one. The only parameters left to Maxim Likelihood Estimation (MLE) thus will be the one regarding market prices of risk, that is the ones involved in the equation

$$\lambda_t = \lambda_0 + \lambda_1 X_t$$

Where  $\lambda_0$  and  $\lambda_1$  will be set at zero initially and after that will be estimated through MLE. The initial values set in the initial step (OLS parameters and market prices of risk) will enter first of all our measurement equation parameters expressions

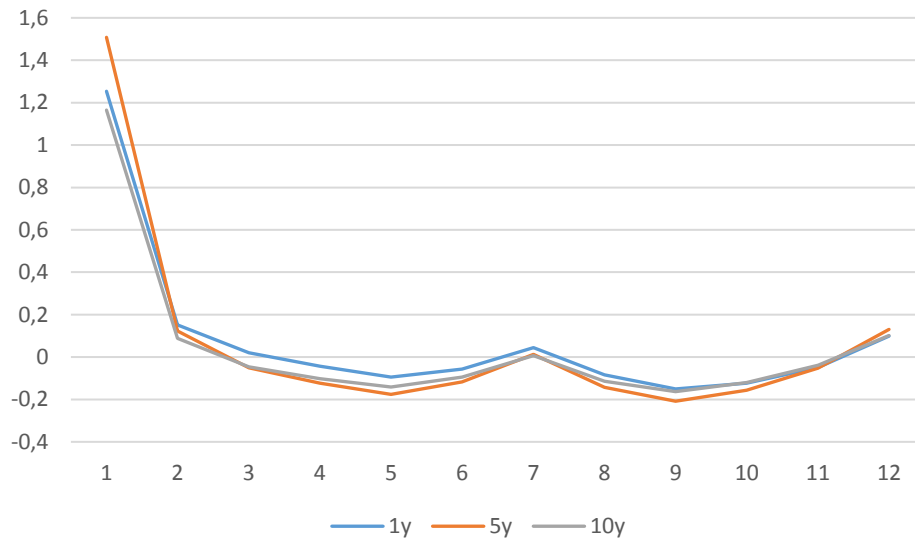
$$\bar{A}_{n+1} = \bar{A}_n + \bar{B}_n'(\mu - \Sigma'\lambda_0) + \frac{1}{2}\bar{B}_n'\Sigma\Sigma'\bar{B}_n - \delta_0$$

$$\bar{B}_{n+1}' = \bar{B}_n'(\Phi - \Sigma'\lambda_1) - \delta_1'$$

As it can be easily seen from the above equations, and as the Ang Piazzesi (2003) work suggests, the measurement equation are recursive processes. With the initial values found in the initial step the first vectors of measurement equation parameters are found (with market prices of risk parameters at zero) and the MLE is performed to find the optimal market prices of risk holding  $A_n$  and  $B_n$  fixed. After this first iteration the parameters are reestimated with the previous optimal estimates for  $\lambda_0$  and  $\lambda_1$  and with new vectors of  $A_n$  and  $B_n$ , and so on until the iteration that maximizes the log likelihood function and that will contain the optimal values for the model. The likelihood function that is used retrieves the distribution of  $Y_t$  from the transformation of yields into a linear transformation of the state variables plus an error, at each round of iteration new measurement equation is computed and this gives back an error  $u_t$ , that is supposed to be distributed normally. In this case the factor entering the measurement equation will only be the 1 month rate, included all the lags, hence the final vector of parameters in the measurement equation will transform the last twelve values for the one month short rate into a

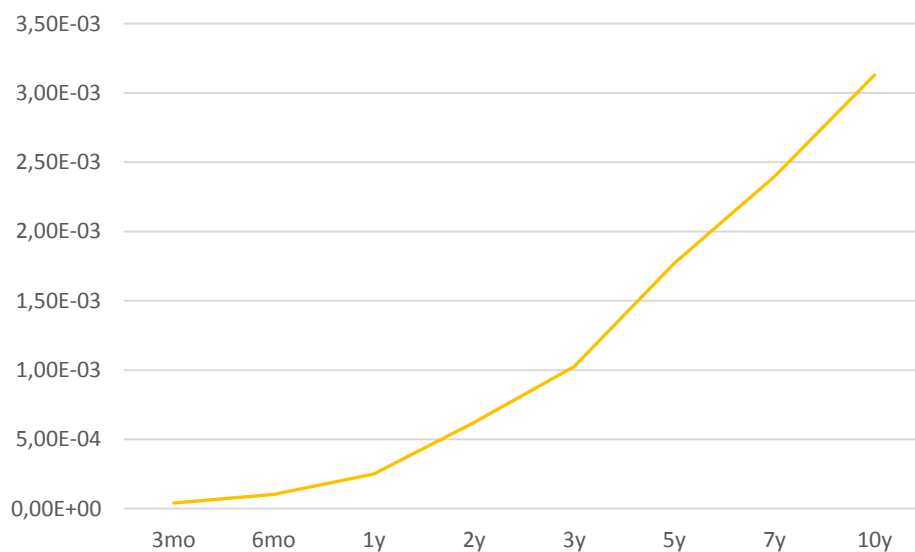


9 maturities vector of yields. In the graph below you can find the effect of the past twelve months on the one year, the five years and the ten years yields respectively.



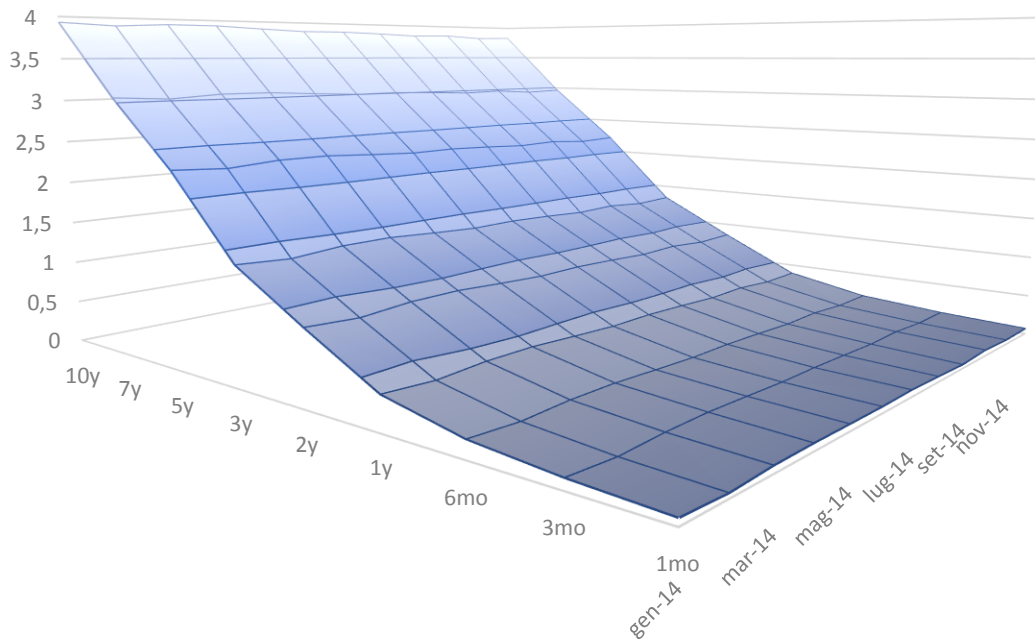
*Figure 4 The effect of Lagged Values of One Month rate on Longer Maturities*

From *Figure 4* we can see that the multipliers for the lagged values of the short rate, given the results of our model are rather similar across maturities. The seventh lagged value for the short rate seems to play an important role together with the first lagged values and the twelfth lag, on the other hand generally we have no prominent differences in the effects across different maturities.



*Figure 5 Intercept of Measurement Equation*

In *Figure 5* we can find the intercept values  $A_n$  that optimize our model. As we expected, for longer maturities the intercept is higher since the lagged values for short term yields will affect the long term yields for a small part, and the mean is higher than the other maturities.



*Figure 6 Model One - Yield Curve Forecast over 2014*

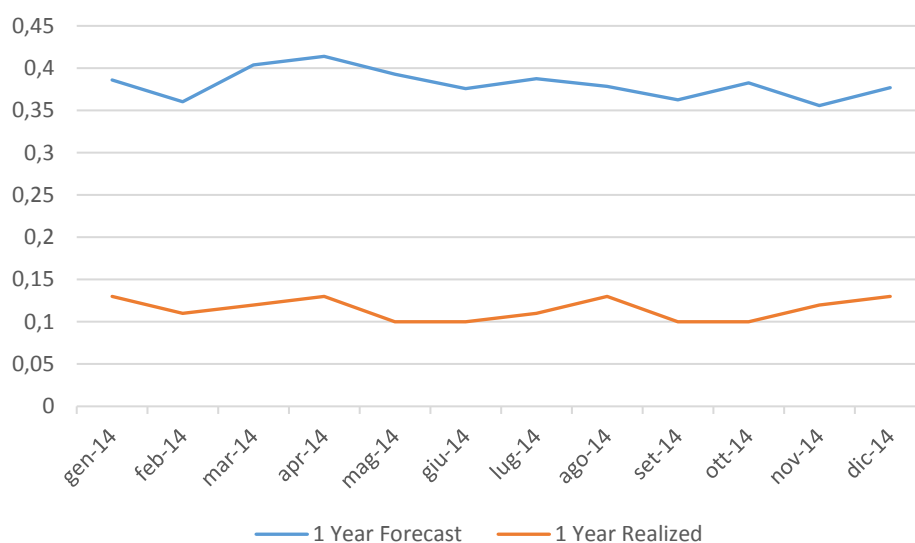
In *Figure 6* the final forecast from the model is presented. Here both the time dimension and the maturity dimension are reported so that a double view could be exploited: yields curve for given forecast periods and evolutions over 2014 of the different maturities. Of course a similar graph is not explicative of our results but surely makes clear what the base for model comparison is and what we are forecasting at this stage. To gain a somewhat practical view on our result we performed a Root Mean Squared Error of the forecasts in order to have an easier method for comparison with the other models.

Maturity	RMSE_Mod1
1 Month	0,0674
3 Months	0,1139
6 Months	0,1924
1 Year	0,3817
2 Years	0,8466
3 Years	1,3390
5 Years	2,2495
7 Years	3,0045
10 Years	3,8758

*Table 1: Root Mean Squared Error on Forecasts over 2014 Period for the entire Yield Curve*

For the root mean squared error we may observe that much of the error is found on longer maturities. This is of course because the model we are estimating is basing on the additional risk premia per period of maturity, thus the error that will produce on the single period will be at its highest when the periods of maturities are much more, thus with longer maturities.

For a further idea of the result of our forecast, we focus on the yield showing the greatest standard deviation over our sample, i.e. the one-year yield.



*Figure 7 1 Year Maturity Yield - Forecast vs. Realized Data*

The main problem with this first estimation is that including crisis period from 2001 on (considering also the Dot-Com bubble) the level of the 1 Year yield is substantially higher for 2014 forecast than realized data. The model thus does not perform very well on this basis, since the volatility of the 2001-2014 period crucially biases forecasts. Even knowing this we will

anyway keep the whole sample together because it is at the basis of our analysis to challenge this kind of models on a dataset that requires a strong model because of such volatility of the market.

In this second model we are going to change the base set of variables. The great part of interest rates finance literature focused on models like this one, where the setting was kept unchanged and equal to affine term structure models, but the state vector was changed. One example is the work of Litterman and Scheinkman (1991)<sup>13</sup>, where the factors affecting the curve are unobservable variables that describe the level, slope and curvature of the yield curve. Beyond this example many more models were created with different kinds of unobservables that described the common movement of yields across maturities, the approach we will employ in this context is the one by Chen and Scott (1993)<sup>14</sup>. Such approach consists in a Kalman Filter applied to interest rate yields in a two step process, where the first step extracts some unobservable factors, while the second one produces estimates of parameters including in the state vector also the unobservables.

Thus going through the first step we have to select three yields that will be the base for our unobservable factors: in our models these yields will be indicated as yields measured *without error*, while the other yields will be indicated as yields measured *with error*. The estimation procedure is rather tricky and not so easy to get conceptually. After some initial arbitrary guesses for model parameters, and thus parameters included in the equations

$$X_t = \mu + \Phi X_{t-1} + \Sigma e_t$$

$$r_t = \delta_0 + \delta_{11}X_t^o + \delta_{12}X_t^u$$

$$\lambda_t = \lambda_0 + \lambda_1 X_t$$

We can get to some initial values for the measurement equation parameters, i.e.

<sup>13</sup> Litterman, R., Scheinkman, J., 1991. Common factors affecting bond returns. *Journal of Fixed Income* 1, 51–61.

<sup>14</sup> Chen, R.R., Scott, L., 1993. Maximum likelihood estimation for a multi-factor equilibrium model of the term structure of interest rates. *Journal of Fixed Income* 3, 14–31.

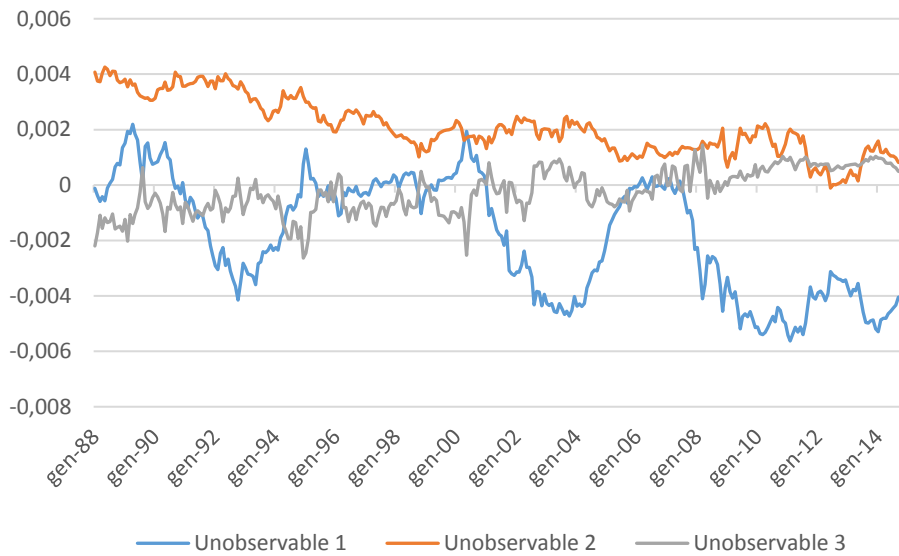
$$\bar{A}_{n+1} = \bar{A}_n + \bar{B}'_n(\mu - \Sigma'\lambda_0) + \frac{1}{2}\bar{B}'_n\Sigma\Sigma'\bar{B}'_n - \delta_0$$

$$\bar{B}'_{n+1} = \bar{B}'_n(\Phi - \Sigma'\lambda_1) - \delta'_1$$

These parameters then, can be used to retrieve from the measurement equation

$$y_t^n = A_n + B'_n X_t$$

Operating such inversion we get to the extraction of the unobservable variables directly from the model through maximization, using subsequent iterations. This lets us to construct new state unobservable variables, already accounting for the market prices of risk and the no arbitrage assumptions included in the model. The behaviour of such unobservables from one month, two years and ten years maturities along the sample period is graphed in *Figure 8*.



*Figure 8 Unobservable Variables extracted from 1 Month, 2 Years and 10 Years Yields*

The choice made for yields without error has been caused by the fact that we wanted to have three unobservables extracted from three different maturities, in such a way that we had some degree of description of the short, medium and long end of the yield curve. Once the unobservables are determined at inception, the model estimates the Maximum Likelihood function based on such variables, the whole thing is then repeated at each iteration until the model is maximized with respect to the vector of parameters of the entire model. Such procedure is totally different from the one taken in model one, but it is perfectly in line with the affine models based on unobservable variables. What we obtain at the end, as in Model 1, are

estimates for the parameters in the measurement equation  $A_n$  and  $B_n$ . In this case these parameters represent the diverse reactions of different maturities to movements in the unobservable factors.

As for Ang and Piazzesi (2003), we can trace in these three unobservable, represented in *Figure 9*, factors the relationships among parameters across different maturities: moving for example the first latent factor will affect mostly the central portion and the short end of the curve, thus it may be assimilated to the “*steepness*” factor in Litterman and Scheinkman (1991). The second unobservable instead affects in a similar fashion all the maturities up to 120 months (10 years), and ensembles the “*level*” factor in literature, while the third unobservable moves the short end and nearly does not affect the medium and long portion, hence it may represent the “*curvature*” effect.

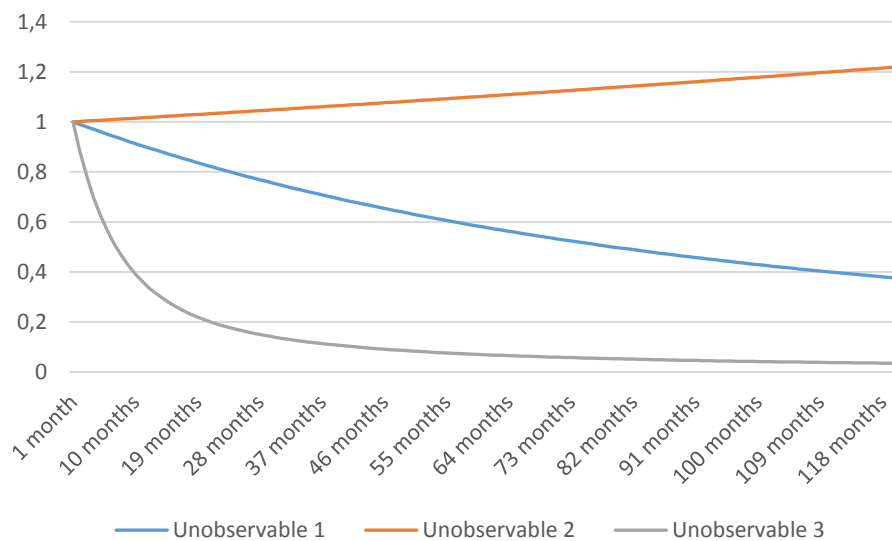
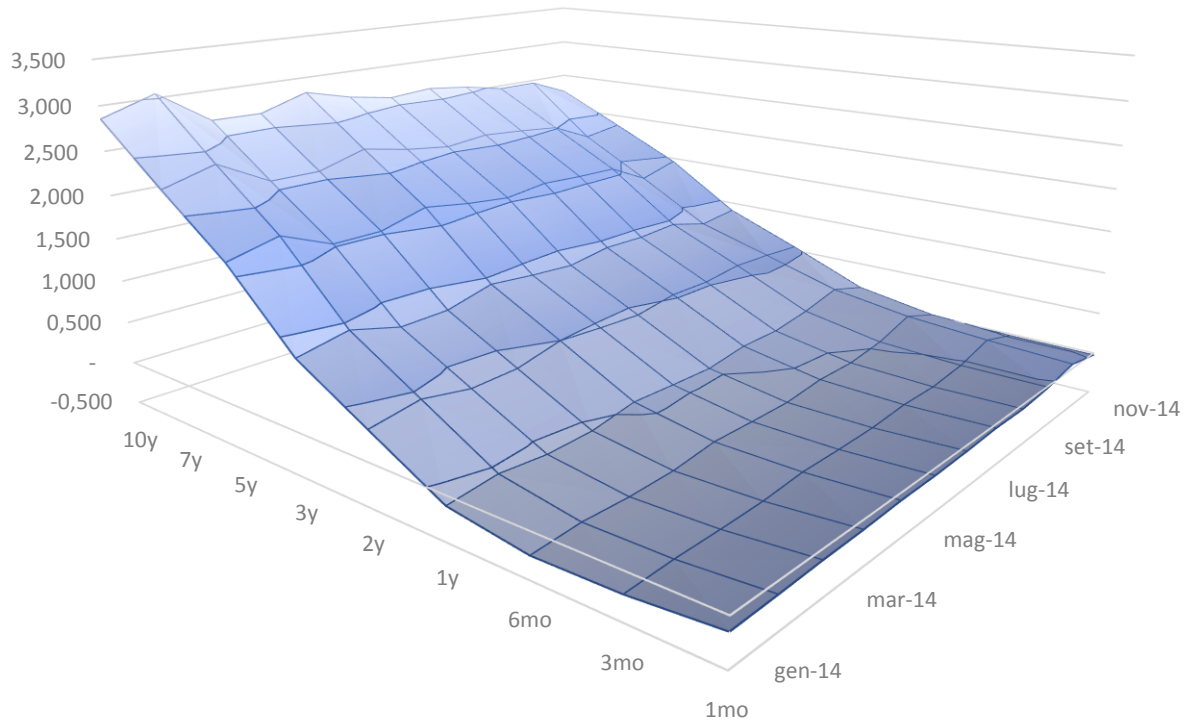


Figure 9 Measurement Equation Parameters estimates for Maturities up to 120 months

From these parameters we obtain the forecasts for 2014 for the initial nine yields we analysed. The forecast is, as in the previous model, obtained estimating the model until time  $t$  and applying it to forecast the  $t+1$  yield curve. In *Figure 10* we observe that the data obtained show negative values for shortest rates at the beginning of the year, with an upward trend in the last

months. For longer maturities we observe no significant trend, but a rather constant pattern of yields.



*Figure 10 Model Two - Forecast for Yield Curves over 2014*

Given this result, an interesting focus has to be made on such short negative yields. In *Figure 11* we have the maturities until one year detailed and compared with realized data in 2014. The forecasted one month rate, measured without error, is forecasted with a dramatic trend in the course of 2014, from  $-0,2$  to almost 0 in twelve months. In the actual data instead the short rate experienced a rather stable pattern, remaining always above 0. Negative rates may be explained as the global trend of rates in the long term is decreasing, but in real world having negative short rates is not that conventional, and our model does not take into account this.



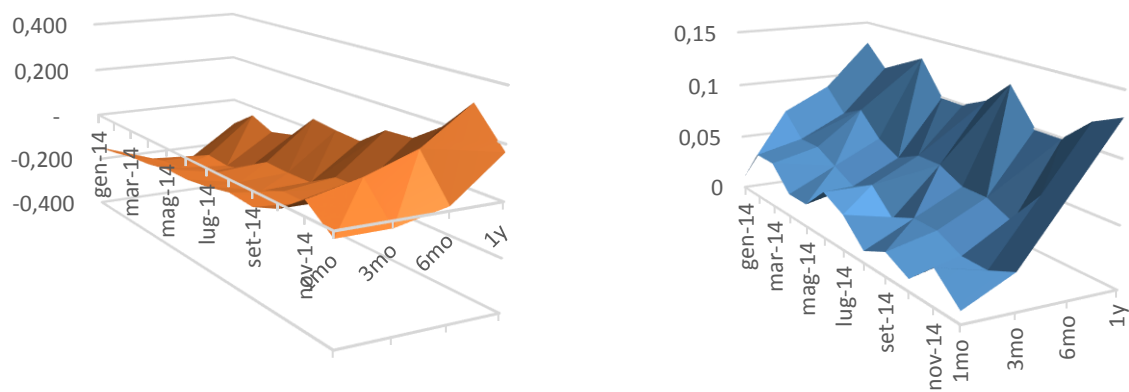


Figure 11 Forecast vs Realized Data for Short end of the Curves over 2014

Having such a hint of errors produced by the model and passing to quantitative analysis, the Root Mean Squared Error can give us a more precise.

Maturity	RMSE_Mod1	RMSE_Mod2
1 Month	0,0674	0,2193
3 Months	0,1139	0,1062
6 Months	0,1924	0,1564
1 year	0,3817	0,3363
2 years	0,8466	0,7012
3 years	1,3390	1,0468
5 years	2,2495	1,6749
7 years	3,0045	2,2277
10 years	3,8758	2,9315

Table 2 Root Mean Squared Error on Forecasts over 2014 Period for the entire Yield Curve

The performance for this second model proves to be slightly better than for the first one, as the root mean squared error decreases. Considering the fact that Model 2 produces negative forecasts by the way would not make it so preferable to Model 1. Certainly this kind of model is a more sophisticated one, but still could be improved for example taking into consideration the different combinations of yields to be measured with and without error: another criteria to be used for example to choose the yields without error may be to choose the yields with higher volatilities, that would thus be the most difficult to forecast. By the way our interest to this kind of model is limited since we will use it as a base for comparison with the macro model, that is *Model 3*.

As for the model in Ang and Piazzesi (2003) we want to unify the no arbitrage affine term structure model, as the ones produced in the two sections above, with the traditional VAR approach, that bases the study of the yield curve on the regression of different yields on macroeconomic variables. In this way we will have the complete yield curve (120 predicted maturities in our case) in a model that also considers macroeconomic variables.

Mainly for data availability issues, the chosen variables are *Consume Price Index (CPI)* and *Industrial Production Index (IPI)*. These two variables in fact can well represent the economy state in the sample period taken under analysis, and most of all, they are computed at a monthly frequency. The CPI rate of change can be considered as a good proxy for inflation, the index in fact contains the level of prices for all consumer goods in the economy, while the Industrial Production Index is well representing GDP movements, as one of the interpretations of Gross Domestic Product is usually the sum of all the values produced in one country.

From the empirical side then, the model is a kind of mixture of the first two models with the addition of macro variables. Model 3 in fact is based on the one month rate dynamics, but with the Chen and Scott (1993) method. Thus we will base the estimation of the whole yield curve on the unobservable variable constructed on the one month maturity yield, plus the two differenced series of the macroeconomic variables. Starting from the beginning, the process is the same described in model two, but here the difference is that the extraction of the unobservable latent factor is done including macro factors in the measurement equation. On a stepwise basis, we start from initial guesses for maximum likelihood, of the fundamental parameters of the model include in the equations

$$X_t = \mu + \Phi X_{t-1} + \Sigma e_t$$

$$r_t = \delta_0 + \delta_{11}X_t^o + \delta_{12}X_t^u$$

$$\lambda_t = \lambda_0 + \lambda_1 X_t$$

Peculiarly for this model only, the short rate dynamics will have a valorised  $X_t^o$  vector, made up by the macroeconomic variables. Once the initial guesses are set, 120 parameters for the measurement equations will be constructed, one for each maturity, and will of course contain at this round also parameters for macroeconomic variables that will enter in

$$\bar{A}_{n+1} = \bar{A}_n + \bar{B}_n'(\mu - \Sigma'\lambda_0) + \frac{1}{2}\bar{B}_n'\Sigma\Sigma'\bar{B}_n - \delta_0$$

$$\bar{B}_{n+1}' = \bar{B}_n'(\Phi - \Sigma'\lambda_1) - \delta_1'$$

These new parameters then will be employed to inverse the measurement equation below equation, where the observable part of  $X_t$  will contain the macroeconomic variables while the unobservable part will be obtain from inversion.

$$y_t^n = A_n + B_n^{o'}X_t^o + B_n^{u'}X_t^u$$

Then again as in Model 2 the maximum likelihood will be maximized based on both observables and unobservables, obtaining the optimal values for the vector of parameters. At the following iteration this optimal values will be the initial values, until the maximum likelihood is maximized.

It seems thus crucial to insert good values for macroeconomic variables. Even having nominal rates as values for yields, our intent is to have a good description of relationships among shocks in the variables. We thus prefer to work with first differenced values for macro variables, resulting in a series as the one represented below

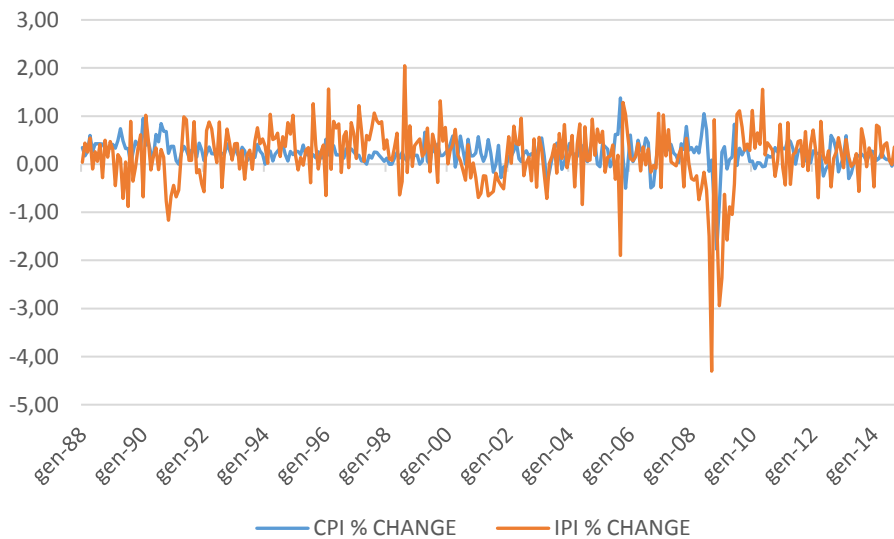
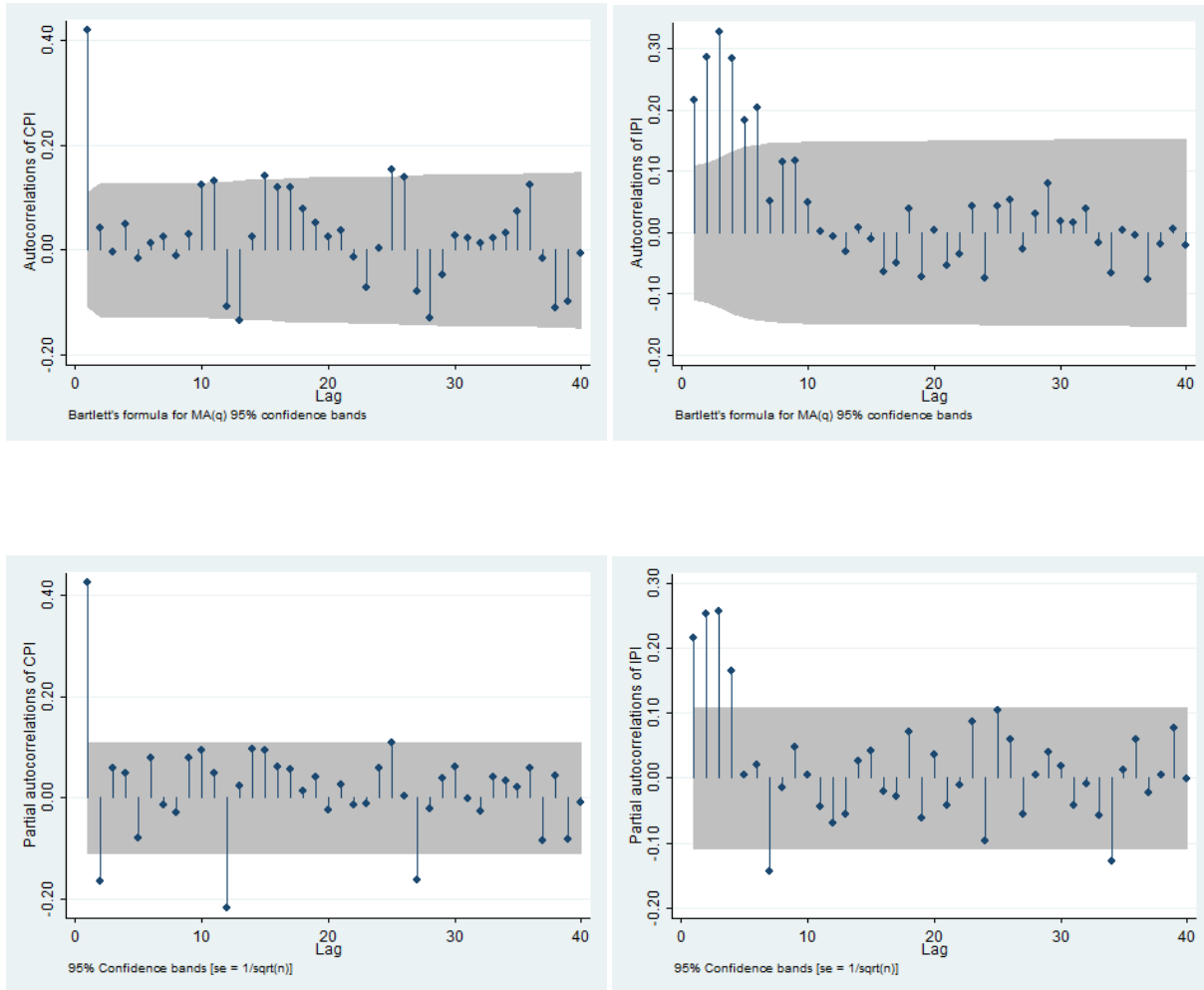


Figure 12 Macroeconomic Variables Time Series in First Differences

Moreover we conducted some test for stationarity and represent autocorrelation functions and partial autocorrelation functions on both the series and the result is the one reported in *Figure 13*.



*Figure 13 Autocorrelation and Partial Autocorrelation Functions for Macroeconomic Variables*

Of course we may not admit stationarity of the series but surely the persistence of the data has been decreased from the level variables and thus should report better estimates. Significant Autocorrelations and partial autocorrelations are few, we thus could treat these variables as stationary.

From the model thus we will come up with three variables, one representing the unobservable part, the other two equal to the macroeconomic variables, and we will have coefficient for all of them, summed up in the measurement equation parameters.

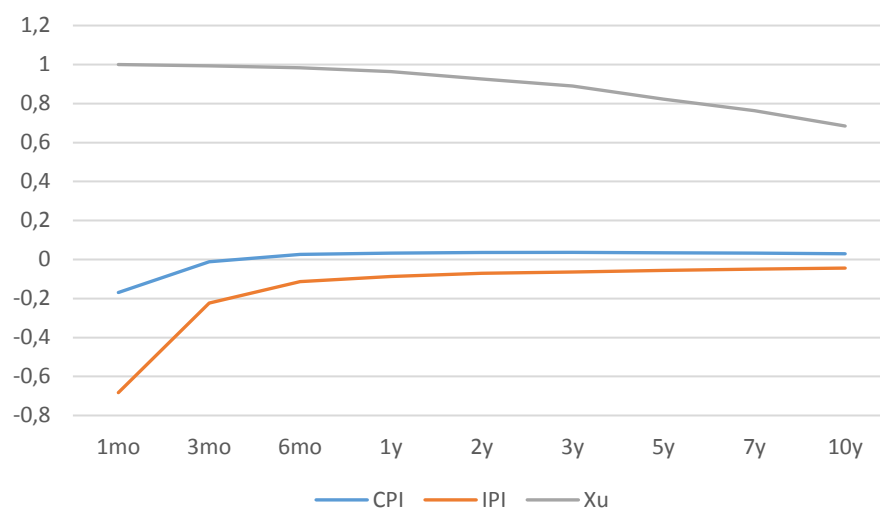


Figure 14 Measurement Equation Parameters for the Considered Maturities

The relationship between factors and different yields shows that Consumer Price Index has negative effects on the shortest maturities, but a positive effect on longer ones, this may seem to go against the Taylor Rule. For the Industrial production Index instead the relationship is always negative even following the shape of the Consumer Price Index, this would suggest that an increase in production would lower interest rates along the whole yield curve. These effects do not appear to be as expected, but this may be due mainly to high volatilities during the crisis period.

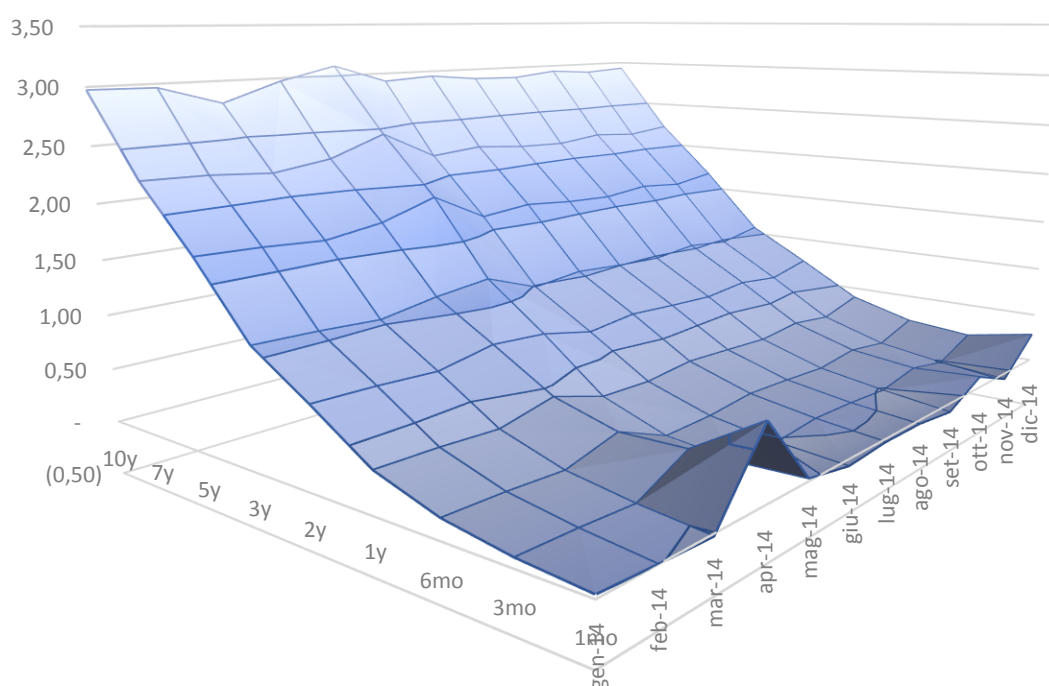


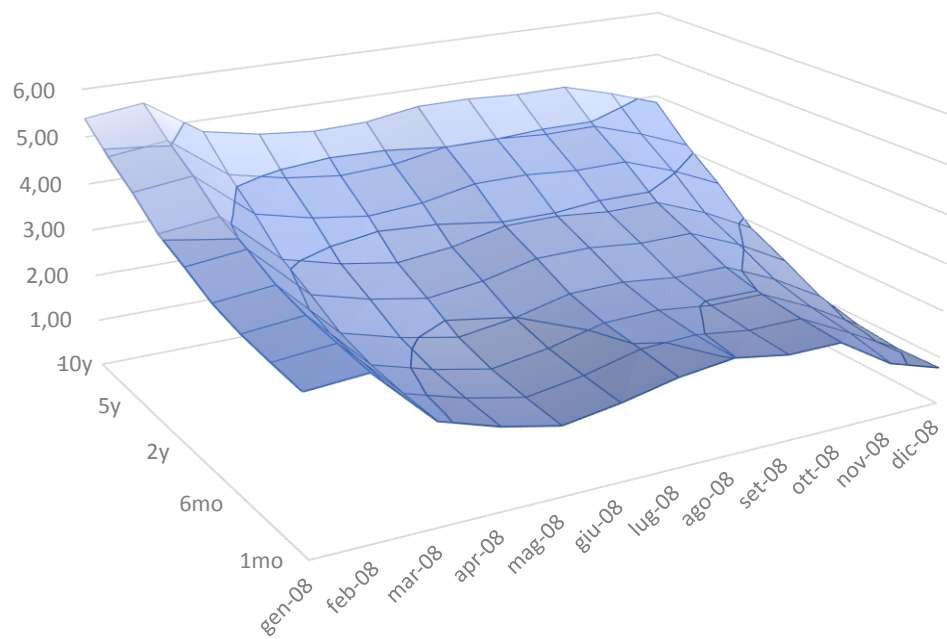
Figure 15 Model 3 - The Forecast of the Yield Curve over 2014

With respect to the previous models, Model 3 show a substantial increase in precision for one step ahead forecast. All of the curves for 2014 in fact are in line, at least for a scale matter, with the realized values for the curve. The inclusion of macroeconomic variables in fact does include important information for interest rates, both for the short and the long end. The short end of the curve is the portion that the central bank can govern somehow through Federal Fund Rate setting, while the long end is usually made up of expectation of investors. Both of these components seem to be well described by this last model: the macro variables included are good proxies for what affects the short maturities, and thus FED policies, and they also are enough correlated with the expectations affecting the long end of the curve.

Maturity	RMSE_Mod1	RMSE_Mod2	RMSE_Mod3
1 Month	0,0674	0,2193	0,0880
3 Months	0,1139	0,1062	0,1368
6 Months	0,1924	0,1564	0,1453
1 year	0,3817	0,3363	0,1329
2 years	0,8466	0,7012	0,4808
3 years	1,3390	1,0468	0,8836
5 years	2,2495	1,6749	1,5669
7 years	3,0045	2,2277	2,0874
10 years	3,8758	2,9315	2,6383

*Figure 16 The Three Models RMSE*

A last interesting point to make is the capacity of prediction in crisis conditions. We run the model on the 2008 period to verify if the model even having parameters based on the whole sample was able to estimate the breaks for the Subprime crisis. The result we had was fairly good since the yields downward shocks (particularly in February and September 2008) have been reported by the model even with a few lag delate.



*Figure 17 The Forecasted Yield Curve over 2008 Period*





Generally thus the model we used seems to well predict the one step ahead values for the interest rates. This may seem not a great result for the limited horizon of the forecast but thinking about the fact that the maturities we estimated are 120 there would be a potentially wide space for hedging or trading strategies linked to such a model. Even in a crisis context in fact the one month ahead forecast are good ones and repeat the shocks suffered by the yield curve even few periods after the shock. Ang and Piazzesi (2003) hence seem to be confirmed as a final conclusion: macro variables do improve finance term structure models mainly through out of sample forecasting.

The analysis may be extended in many more ways, for example further unobservables or different macro variables may be included after a preliminary analysis on the single factor. Active portfolio strategies may be constructed and components including structural breaks may be considered.



- J.C. Cox, J.E. Ingersoll, S.A. Ross, 1981, “The relation between forward prices and future prices”, *Journal of Financial Economics* 9, 321-346
- M. Fisher, 1998, “A Simple Model of the Failure of the Expectations Hypothesis”, Working paper (Federal Reserve Bank of Atlanta, Atlanta, GE)
- Vasicek, O., 1977. An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics*, 5, 177-88.
- Refet S. Gürkaynak Brian Sack and Jonathan H. Wright, “The U.S. Treasury Yield Curve: 1961 to the Present”, 2006.
- Taylor, J. B., 1993. Discretion versus Policy Rules in Practice, *Carnegie-Rochester Conference Series on Public Policy*, 39, 195-214.
- Clarida, R., Gali, J., Gertler, M., 2000. Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory. *Quarterly Journal of Economics*, 65, 147-80.
- Campbell J. Y., 1995, Some Lessons from the Yield Curve, *The Journal of Economic Perspectives*, Vol. 9, n. 3, 129-152.
- Ang, Andrew and Monika Piazzesi (2003) “A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables.” *Journal of Monetary Economics* 50, 745–787.
- Dai, Q., Singleton, K., 2000. Specification analysis of affine term structure models. *Journal of Finance* 55, 1943–1978
- Harvey, C.R., 1988. The real term structure and consumption growth. *Journal of Financial Economics* 22, 305–333.
- Estrella, A., Hardouvelis, G.A., 1991. The term structure as a predictor of real economic activity. *Journal of Finance* 46, 555–576.
- Litterman, R., Scheinkman, J., 1991. Common factors affecting bond returns. *Journal of Fixed Income* 1, 51–61.
- Chen, R.R., Scott, L., 1993. Maximum likelihood estimation for a multi-factor equilibrium model of the term structure of interest rates. *Journal of Fixed Income* 3, 14–31.

Duffee, G. R., 2002. Term Premia and Interest Rate Forecasts in Affine Models, *Journal of Finance*, 57, 405-43.

Duffie, D., Kan, R., 1996. A Yield-Factor Model of Interest Rates, *Mathematical Finance*, 6, 379-406.

Constantinides, G. M., 1992. A Theory of the Nominal Term Structure of Interest Rates, *Review of Financial Studies*, 5, 531-52.

Evans, C. L., Marshall, D. A., 1998. Monetary Policy and the Term Structure of Nominal Interest Rates: Evidence and Theory, *Carnegie-Rochester Conference Series on Public Policy*, 49, 53-111.

#### **Websites**

<http://www.bea.gov/>

<http://www.usinflationcalculator.com/>

[www.fred.stlouisfed.org](http://www.fred.stlouisfed.org)

[www.federalreserve.gov/releases/h15/current/h15.pdf](http://www.federalreserve.gov/releases/h15/current/h15.pdf)





## SPECIAL THANKS

---

I want to thank first of all professor Nunzio Cappuccio, that supported me all along the period of work on the thesis and that strongly encouraged and believed in me. A great thank to my mother Ofelia that sustained and worked hard to allow me to get my own path to graduation. A great thank to Chiara that still sustains me with her love and help. Thanks to my brothers Roberto, Federico, Alessio and Giuliano that are the cornerstone of my strength and commitment. Thanks to my big brothers Chico, Enrico and Marco for being with me always and anyway. Thanks to my great friends and flatmates Stefano, Edoardo and Marco, to my Erasmus friends, to Duckbill Roll, to all the colleagues in Hollywood, and to everyone in this great big family.

Thank you.

*A chi mi ha aiutato, mentre mi arrangiavo*